

# THE OECD ANALYTICAL BERD (ANBERD) DATABASE

OECD/DSTI/EAS

# LA BASE DE DONNÉES ANALYTIQUE SUR LA DIRDE (ANBERD) DE L'OCDE

OCDE/DSTI/EAS

## Contents

INTRODUCTION.....	2
THE OECD ANALYTICAL BERD (ANBERD) DATABASE: EXPLANATIONS IN ITS DEVELOPMENT .....	4
Limitations of the official BERD (OFFBERD) data .....	4
The OECD ANalytical BERD (ANBERD) database .....	5
What's new since the 2009 ANBERD edition? .....	6
ANBERD ESTIMATES.....	6
Historical data .....	6
General estimation methods in ANBERD .....	7
Special estimation method for the United States .....	7
Product field data .....	8

## Table des Matières

INTRODUCTION.....	9
LA BASE DE DONNÉES ANALYTIQUE SUR LA DIRDE (ANBERD) DE L'OCDE : EXPLICATIONS CONCERNANT SON ÉLABORATION .....	11
Les limites des données officielles sur la DIRDE (OFFBERD) .....	11
La base de données analytique de l'OCDE sur la DIRDE (ANBERD)....	12
Nouveautés depuis l'édition 2009.....	13
ESTIMATIONS ANBERD .....	13
Données historiques .....	14
Méthodes générales d'estimation dans ANBERD .....	14
Méthode d'estimation ANBERD pour les Etats-Unis .....	14
Données par groupes de produits .....	15

## INTRODUCTION

Industrial research & development (R&D) is defined as R&D activities carried out in the business enterprise sector, regardless of the origin of funding. While the government and the higher education sectors also carry out R&D activities, industrial R&D remains the most closely linked to the creation of new products and production techniques.

The ANalytical Business Enterprise Research and Development database (ANBERD) was developed to provide analysts with comprehensive and internationally comparable data on industrial R&D expenditures that address the problems of international comparability and breaks in the time series of official business enterprise R&D data (OFFBERD) provided to the OECD by its member countries through the joint OECD/Eurostat R&D survey. The ANBERD database includes a number of estimations and is published under the responsibility of the Secretary General of the OECD as it does not represent Member countries' official submissions of business enterprise R&D data.

The ANBERD database presents industrial R&D expenditure data broken down in 60 manufacturing and services sectors (Table 1 shows the industry list used in ANBERD) for OECD countries and selected non-member economies. The reported data are in [ISIC Revision 3.1](#), start from 1987 and are expressed in national currencies as well as in PPP US dollars, both at current and constant prices. The data are updated continuously on a rolling basis in order to release the tables as new data are collected.

Countries covered include OECD member countries (Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States) and selected non-member economies (China, Romania, the Russian Federation, Singapore, South Africa, and Chinese Taipei).

ANBERD was conceived so that business R&D expenditure by industry data could be better matched with industrial performance data presented in the STructural ANalysis database, which includes measures of output, labour input and investment by industry ([STAN-industry](#)); the Bilateral Trade in Goods by Industry and End-use Category ([BTDIxE](#)); OECD harmonised Input-Output tables ([STAN-IO](#)); and [STAN Indicators](#). The presence of ANBERD in this "STAN family of databases" allows researchers, for example, to measure R&D intensity (R&D/Output) by industry and undertake analyses of international technology diffusion. The databases are united by compatible industry lists based on ISIC Revision 3.

Further information on the OECD's ANBERD database can be found at [www.oecd.org/sti/anberd](http://www.oecd.org/sti/anberd). ANBERD tables can be accessed via the OECD's data dissemination service [OECD.STAT](#).

The recommended source for use of the ANBERD estimates is "OECD, ANBERD Database".

**TABLE 1. BERD ISIC, Revision 3.1 Classification**

<b>ANBERD code</b>	<b>ISIC Rev.3.1</b>	<b>Description</b>
CTOTAL	01-99	TOTAL BUSINESS ENTERPRISE
<b>C01T05</b>	<b>01-05</b>	<b>AGRICULTURE, HUNTING, FORESTRY AND FISHING</b>
<b>C10T14</b>	<b>10-14</b>	<b>MINING AND QUARRYING</b>
<b>C15T37</b>	<b>15-37</b>	<b>TOTAL MANUFACTURING</b>
C15T16	15-16	Food products, beverages and tobacco
C15	15	...Food products and beverages
C16	16	...Tobacco products
C17T19	17-19	Textiles, textile products, leather and footwear
C17	17	...Textiles
C18	18	...Wearing apparel, dressing and dyeing of fur
C19	19	...Leather, leather products and footwear
C20T22	20-22	Wood, paper, printing, publishing
C20	20	...Wood and products of wood and cork
C21T22	21-22	...Pulp, paper, paper products, printing and publishing
C21	21	.....Pulp, paper and paper products
C22	22	.....Printing and publishing
C23T25	23-25	Chemical, rubber, plastics and fuel products
C23	23	...Coke, refined petroleum products and nuclear fuel
C24	24	...Chemicals and chemical products
C24X	24 less 2423	.....Chemicals excluding pharmaceuticals
C2423	2423	.....Pharmaceuticals
C25	25	...Rubber and plastics products
C26	26	Other non-metallic mineral products
C27	27	Basic metals
C271T31	271+2731	...Iron and steel
C272T32	272+2732	...Non-ferrous metals
C28	28	Fabricated metal products, except machinery and equipment
C29T35	29-35	Machinery and equipment, instruments and transport equipment
C29	29	...Machinery and equipment, nec
C30	30	...Office, accounting and computing machinery
C31	31	...Electrical machinery and apparatus, nec
C32	32	...Radio, television and communication equipment
C321	321	.....Electronic valves and tubes and other electronic components
C32X	32 less 321	.....Television, radio and communication equipment nec
C33	33	...Medical, precision and optical instruments
C34	34	...Motor vehicles, trailers and semi-trailers
C35	35	...Other transport equipment
C351	351	.....Building and repairing of ships and boats
C353	353	.....Aircraft and spacecraft
C352A9	352+359	.....Railroad equipment and transport equipment nec
C36T37	36-37	Manufacturing n.e.c. and recycling
C36	36	...Furniture; manufacturing nec
C361	361	.....Furniture
C369	369	.....Manufacturing nec
C37	37	...Recycling
<b>C40T41</b>	<b>40-41</b>	<b>ELECTRICITY, GAS AND WATER SUPPLY</b>
<b>C45</b>	<b>45</b>	<b>CONSTRUCTION</b>
<b>C50T99</b>	<b>50-99</b>	<b>TOTAL SERVICES</b>
C50T52	50-52	Wholesale and retail trade; repairs
C55	55	Hotels and restaurants
C60T64	60-64	Transport, storage and communications
C642	642	...Telecommunications
C65T67	65-67	Financial intermediation
C70T74	70-74	Real estate, renting and business activities
C72	72	...Computer and related activities
C722	722	.....Software consultancy and supply
C72X	72 less 722	.....Other computer services nec
C73	73	...Research and development
C74	74	...Other business activities
C75T99	75-99	Community, social and personal services

## THE OECD ANALYTICAL BERD (ANBERD) DATABASE: EXPLANATIONS IN ITS DEVELOPMENT

### Limitations of the official BERD (OFFBERD) data

The *Frascati Manual* sets out guidelines for the collection of internationally comparable R&D data, including for the business enterprise sector (BERD). While there is broad adherence to *Frascati* definitions and methodologies throughout the OECD, there are nevertheless significant divergences across countries in the way in which BERD data are collected. These have important implications for international comparability and for time series consistency. The purpose of this section is to provide a general overview of the major difficulties of international comparability associated with the official business enterprise R&D data.

BERD survey sampling methodologies or coverage ratios may differ from one country to another. In some cases, substantial numbers of small firms are not included in the frame of the major survey. Additional surveys of excluded firms may be undertaken, and appropriate estimations made, but this is not a universal practice, and imputation is not done in a consistent way across countries. Treatment of borderline research institutions also varies from country to country. Where government or semi-government agencies and research bodies conduct a substantial amount of industrial R&D, these differences may be important.

The *Frascati Manual* practice is to report BERD on an enterprise basis. When this is interpreted strictly, all the BERD of a diversified enterprise will be allocated to the industrial class of its principal activity. In circumstances where a few large firms dominate R&D spending in several areas, this can and does lead to underestimates of R&D associated with the secondary activities of the firms. R&D is thus overestimated for some industries and underestimated for others. However, not all countries follow a strict enterprise basis for allocating R&D expenditures to industrial classes. Some countries make a disaggregation of the R&D of their largest, diversified firms into a number of different activities, or collect R&D data on a product field basis.

The level of international comparability between countries reporting on a product field or semi-product field basis is likely to be greater than that between those following a strict enterprise approach. One of the advantages of product field data is that the R&D expenditure performed by an enterprise that is active in more than one industry is distributed to the relevant industries.

There are also important differences between countries in the treatment of R&D undertaken by firms normally in the service sector but closely associated (not necessarily contractually) to manufacturing firms. Industrial research institutes, largely funded by the manufacturing industries they serve, are the most frequent examples. The latest *Frascati Manual* (2002) states that the R&D undertaken by such institutes should be attributed to the "Research and development" industry (ISIC Rev. 3.1: 73) but that "additional information should be collected for analytical purposes and international comparisons [...] by requesting product field data". In addition, the R&D expenditures of large public business enterprises such as telecommunications and energy authorities are frequently classified in the service sector even though they undertake considerable, and in some cases primarily, manufacturing R&D. This can distort international comparisons at the aggregate as well as the individual industry level.

Revisions in classification schemes are also common and may lead to significant changes in the concordance between national industrial classifications and ISIC and consequent breaks in the corresponding BERD time series. This limits international comparability of BERD data. In addition, some countries do not conduct a business R&D survey every year, therefore the time coverage of BERD data is uneven. Because BERD is both rising at different rates across OECD countries and highly volatile within countries, considerable caution has to be exercised in making international comparisons based on figures from different years, even when spanning only a few years.

Finally, data may not be available for all industries in a particular country. Frequently, for confidentiality or other reasons, data for individual industries may not be available separately but are instead included with data for other industries. Such treatment is often not consistent over time, so that in some years the industries in question will be available separately and in others they will be suppressed or combined with other data. Time series may thus have large discontinuities, making international comparisons impossible.

As a consequence, the official BERD (OFFBERD) data consist of a fragmented time series that must be used with caution.

### **The OECD ANalytical BERD (ANBERD) database**

Even with a detailed set of explanatory notes, the uneven coverage and large number of discontinuities in the OFFBERD data limit their analytical usefulness. In order to address the inconsistencies associated with the OFFBERD data, the ANBERD database was developed. ANBERD is designed to provide analysts with a comprehensive and internationally comparable data set on industrial R&D expenditures. Nevertheless, by its nature ANBERD contains many OECD estimates which may differ significantly from the corresponding official data. ANBERD is subject to revisions because it depends upon a number of estimations which are constantly being refined and reviewed.

ANBERD is based on OFFBERD but involves estimates wherever:

- there are significant problems associated with the enterprise basis of the survey or with borderline institution classifications. Wherever possible, ANBERD estimates include adjustments to ensure that the BERD of borderline institutions and public enterprises is allocated to the relevant industries;
- there are important deviations (*e.g.* "Aircraft" is included in "Motor Vehicles") from the standard BERD industrial classification;
- there are significant adjustments required due to incomplete survey coverage;
- there are discontinuities or breaks in series due to a change in industrial classification or survey techniques;
- there are missing data for entire years in cases where surveys are not conducted every year.

The ANBERD data set is constructed in close collaboration with national statistical authorities and derestricted on a country-by-country basis. In all cases, statistical estimation techniques for filling in

missing values or to adjust the existing OFFBERD data are used to obtain improved and more extensive time series. The next section provides a brief description of the estimation techniques used in the development of the ANBERD estimates.

### **What's new since the 2009 ANBERD edition?**

ANBERD tables are now updated on a rolling basis and released country-by-country as new data become available.

In the 2009 edition, "Agriculture" and "Mining" were included in national aggregates (total BERD) even though they were not shown separately in the detailed ANBERD tables. In the current edition, they are now published separately. Additionally, more countries are covered: Estonia and Mexico are now included.

## **ANBERD ESTIMATES**

This section presents the current state of the ANBERD database. The ANBERD database consists of time series starting from 1987 in ISIC, Revision 3.1 for all OECD member countries (Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States) and selected non-member economies (China, Romania, the Russian Federation, Singapore, South Africa, Chinese Taipei).

The ANBERD estimates are given in national currencies and in USD purchasing power parities (PPPs), both at current and constant prices. The ANBERD estimates for the Euro area countries are given in Euro (EUR). Data relating to years prior to the year of Euro Zone accession have been converted from national denomination into EUR denomination by applying the irrevocable EUR/national currency conversion rate. This approach facilitates comparisons within one country over time and ensures that the historical evolution is preserved, however the pre-1999 data of the EMU countries expressed in EUR cannot be added across countries to generate Euro area aggregate or to carry out cross-country comparisons.

GDP price indices and PPPs used to deflate and convert the ANBERD data are available in the [Economic Indicators for MSTI \(Main Science and Technology Indicators\) database](#).

### **Historical data**

The previous (2009) edition of ANBERD is still available (OECD.STAT link: [ANBERD 2009](#)) as well as an archived version in ISIC Revision 2 ([ANBERD rev.2](#)) providing data for 15 OECD countries plus the zone total for the European Union from 1973 to 1997/98.

## **General estimation methods in ANBERD**

The principal problems associated with the official BERD data of the ANBERD countries are:

- Missing data for some industries that are combined with data for other industries or included only in sub-totals.
- Breaks in series due to revisions in the industrial classification. The sudden inclusion or exclusion of industries is quite common. For example, “Office & Computing Machinery” may be included with “Non-Electrical Machinery” for some years, but separately available for others.
- Missing data for entire years due to the fact that some countries conduct R&D surveys only every two or three years. This is the case for some countries including Austria, Greece, Ireland (until 2000), Sweden and the Russian Federation.
- For countries where institutes which undertake R&D for the manufacturing sector have been classified in the service sector, thus giving a misleading picture of industrial R&D.

These problems are treated firstly by obtaining information and estimates from the national statistical authorities of the countries concerned, which allows many of the inconsistencies to be solved (missing industries and breaks in series). Secondly, interpolation methods are used to estimate BERD for missing years where this is necessary.

## **Special estimation method for the United States**

For years prior to 2008, approximately one-quarter to one-third of the individual industry entries for official total intramural BERD of the United States have been suppressed for confidentiality reasons associated with federal government funding. In order to include the United States in international comparisons of R&D performance, the missing industries had to be estimated. For years prior to 2008, ANBERD estimates for the United States are based on the annual National Science Foundation (NSF) publication “Research and Development in Industry”, which contains selected tables of company-funded R&D. These data were based on the annual Survey of Industrial Research and Development conducted by the Bureau of the Census for the NSF. For most industries company-funded R&D is sufficient to estimate total R&D because federal government funding of aggregate U.S. industrial R&D is less than 10%. For ANBERD, in order to estimate total R&D performed in each industry, only the federal component of BERD needs to be estimated for the suppressed industries. The estimated government-funded R&D was then added to the published company-funded R&D, to yield total BERD.

The federal shares of BERD were estimated. For most industries, the federal funding was assumed to be very small, meaning that total BERD for these industries could be estimated with some confidence. This allowed the estimation of other industries or sub-totals by simple addition or subtraction. In other industries, federal funding was estimated by linearly interpolating the federal component of BERD between data points where the breakdown between federal and company components were available. The result is a complete matrix (for all missing manufacturing sectors except “Shipbuilding & Repairing”). “Shipbuilding & Repairing” is included with “Other Transportation Equipment” in the NSF publication.

From 1999, NSF data are published using the North American Industry Classification System (NAICS), replacing the former US Standard Industrial Classification (SIC). Even though these data are still comparable with previous years for Total Enterprises, Total Manufacturing and Total Services, breaks in series may occur at a more disaggregated level of ISIC, Revision 3.1.

Beginning in the late 1990s, increasingly large amounts of R&D were attributed to the wholesale trade industries, although most of this growth was shown to be due to two factors: (1) the payroll-based methodology used to assign industry classifications; and (2) the change to NAICS in 1999. This led to a revision of the method used to classify R&D by industry beginning with the reference year 2004. The major impact of this revision was a 40% increase of the amount of R&D allocated to the manufacturing sector (mainly in pharmaceuticals and medicines, as well as ICT products), mainly at the expense of the wholesale trade industry.

For years after 2007, the Survey of Industrial Research and Development (SIRD) was replaced by the Business R&D and Innovation Survey and ANBERD estimates for the United States are based on the annual National Science Foundation (NSF) publication "Business R&D and Innovation." As with previous years' statistics, some individual industry entries for official total intramural BERD of the United States have been suppressed for confidentiality reasons associated with federal government and other sources of funding. In order to include the United States in international comparisons of R&D performance, the missing industries had to be estimated.

### **Product field data**

Following the recommendations of the latest *Frascati Manual* (2002), one of the aims of ANBERD is to present product field data, whenever possible. Given differences in the classification of firms by industry in national surveys, and despite efforts to harmonise, caution must be exercised when comparing data for specific sectors across countries (see also the country notes under each data table). For example:

- Some countries collect and submit product field data (e.g. Finland, Sweden, United Kingdom);
- For others, adjustments are made using product field data to reduce enterprise bias (e.g. Japan, Norway, Spain) or redistributing the R&D of ISIC R-3 73 (R&D services) by industry served;
- Finally, for the rest, product field data are not available so data are reported by principal activity with some additional estimations by the OECD for missing data (e.g. Canada, United States);

For three countries (Belgium, the Czech Republic and South Africa), data are available using both breakdowns, so two tables are included in ANBERD.

## INTRODUCTION

La Recherche-Développement (R-D) industrielle correspond aux activités de R-D menées dans le secteur des entreprises, quelle que soit leur source de financement. Si le secteur de l'État et celui de l'enseignement supérieur réalisent également des travaux de R-D, le secteur industriel demeure le plus étroitement lié à la création de nouveaux produits et techniques de production.

La base de données analytique des dépenses en recherche et développement dans l'industrie (ANBERD) a été conçue afin de fournir des séries de données cohérentes et comparables à l'échelle internationale. Cet ensemble de données permet également de surmonter les problèmes liés aux ruptures de séries qui affectent les données officielles de dépenses en R-D (OFFBERD) collectées par les pays membres de l'OCDE par le biais de l'enquête de l'OCDE/Eurostat sur la R-D du secteur des entreprises. La base de données ANBERD contient des estimations et est publiée sous la responsabilité du Secrétaire général de l'OCDE puisqu'elle ne représente pas les chiffres officiels soumis par les pays Membres sur les dépenses de R-D du secteur des entreprises.

La base ANBERD présente des données annuelles sur les dépenses de R-D du secteur des entreprises. Les données sont publiées pour l'ensemble des pays de l'OCDE ainsi que pour une sélection de pays non membres et sont ventilées en 60 industries manufacturières et des services. Les séries sont présentées dans la classification [CITI Révision 3.1](#), débutent en 1987 et sont exprimées en monnaie nationale ainsi qu'en parité de pouvoir d'achat (\$ PPA), à la fois aux prix courants et constants. Les données sont publiées de "manière glissante" : tout nouveau tableau collecté est traité et publié dès sa finalisation pour une mise à jour continue, pays par pays.

ANBERD couvre l'ensemble des pays de l'OCDE (l'Australie, l'Autriche, la Belgique, le Canada, le Chili, la République tchèque, le Danemark, l'Estonie, la Finlande, la France, l'Allemagne, la Grèce, la Hongrie, l'Islande, l'Irlande, Israël, l'Italie, le Japon, la Corée, le Luxembourg, le Mexique, les Pays-Bas, la Nouvelle-Zélande, la Norvège, la Pologne, le Portugal, la République slovaque, la Slovénie, l'Espagne, la Suède, la Suisse, la Turquie, le Royaume-Uni et les États-Unis) ainsi qu'une sélection de pays non membres (Afrique du Sud, Chine, Fédération de Russie, Roumanie, Singapour, Taipei chinois).

La base de données ANBERD a été conçue de façon à pouvoir relier les données de dépenses de R-D des entreprises par industrie aux données de performance industrielle présentées dans la base de données STAN (analyse structurelle) qui contient des mesures de production, de main d'œuvre, et d'investissement par industrie ([STAN-industrie](#)) ; la base sur le commerce bilatéral des biens par Industrie et catégorie d'utilisation finale ([BTDIXE](#)) ; les tableaux entrées-sorties harmonisés de l'OCDE ([STAN-IO](#)) et les [indicateurs de STAN](#). La présence de ANBERD dans cette « famille des bases de données STAN » permet aux chercheurs, par exemple, de mesurer l'intensité de R-D (R-D/valeur ajoutée) par industrie et de mener des analyses sur la diffusion internationale des technologies. Les bases de données sont reliées par des listes d'industries compatibles fondées sur la CITI révision 3.

De plus amples informations sur la base ANBERD sont disponibles sur [www.oecd.org/sti/anberd-fr](http://www.oecd.org/sti/anberd-fr). Les données sont accessibles depuis le service de dissémination des données de l'OCDE [OECD.STAT](#).

La source recommandée pour l'utilisation des estimations ANBERD est « OCDE, base de données ANBERD ».

**TABLE 1.** Classification DIRDE CITI, révision 3.1

<b>code ANBERD</b>	<b>CITI Rév.3.1</b>	<b>Description</b>
CTOTAL	01-99	TOTAL ENTREPRISES
<b>C01T05</b>	<b>01-05</b>	<b>AGRICULTURE, CHASSE, SYLVICULTURE ET PÊCHE</b>
<b>C10T14</b>	<b>10-14</b>	<b>ACTIVITÉS EXTRACTIVES</b>
<b>C15T37</b>	<b>15-37</b>	<b>TOTAL INDUSTRIES MANUFACTURIÈRES</b>
C15T16	15-16	Produits alimentaires, boissons et tabac
C15	15	...Produits alimentaires et boissons
C16	16	...Produits à base de tabac
C17T19	17-19	Textiles, articles d'habillement, cuir et chaussures
C17	17	...Textiles
C18	18	...Habillement, préparation et teinture des fourrures
C19	19	...Cuirs, articles en cuir et chaussures
C20T22	20-22	Bois, papier, imprimerie, édition
C20	20	...Production de bois, articles en bois et en liège
C21T22	21-22	...Pâtes et papier, articles en papier, imprimerie et édition
C21	21	.....Pâtes, papier, carton, articles en papier et carton
C22	22	.....Imprimerie et édition
C23T25	23-25	Produits chimiques et pétroliers, caoutchouc et plastique
C23	23	...Cokéfaction, produits pétroliers raffinés et combustibles nucléaires
C24	24	...Produits chimiques
C24X	24 less 2423	.....Produits chimiques, sauf produits pharmaceutiques
C2423	2423	.....Produits pharmaceutiques
C25	25	...Articles en caoutchouc et matières plastiques
C26	26	Autres produits minéraux non métalliques
C27	27	Produits métallurgiques de base
C271T31	271+2731	...Métaux ferreux
C272T32	272+2732	...Métaux non ferreux
C28	28	Fabrication d'ouvrages en métaux (sauf machines et matériel)
C29T35	29-35	Machines et matériel, instruments, transports
C29	29	...Machines et matériel nca
C30	30	...Machines de bureau, comptables et informatiques
C31	31	...Machines et appareils électriques nca
C32	32	...Appareils de radio, télévision et communication
C321	321	.....Tubes, valves et autres composants électroniques
C32X	32 less 321	.....Appareils de radio, télévision et communication nca
C33	33	...Instruments médicaux, de précision, d'optique et d'horlogerie
C34	34	...Véhicules automobiles, remorques et semi-remorques
C35	35	...Autres matériels de transport
C351	351	.....Construction et réparation de navires
C353	353	.....Construction aéronautique et spatiale
C352A9	352+359	.....Matériel ferroviaire roulant ; équipements de transport nca
C36T37	36-37	Industries manufacturières n.c.a. et récupération
C36	36	...Meubles ; activités de fabrication nca
C361	361	.....Meubles
C369	369	.....Activités de fabrication nca
C37	37	...Récupération
<b>C40T41</b>	<b>40-41</b>	<b>ÉLECTRICITÉ, GAZ ET EAU</b>
<b>C45</b>	<b>45</b>	<b>CONSTRUCTION</b>
<b>C50T99</b>	<b>50-99</b>	<b>TOTAL SERVICES</b>
C50T52	50-52	Commerce de gros et de détail ; réparations
C55	55	Hôtels et restaurants
C60T64	60-64	Transports, entreposage et communications
C642	642	...Télécommunications
C65T67	65-67	Intermédiation financière
C70T74	70-74	Immobilier, locations et activités de service aux entreprises
C72	72	...Activités informatiques et activités connexes
C722	722	.....Production de logiciels
C72X	72 less 722	.....Autres activités informatiques
C73	73	...Recherche et développement
C74	74	...Autres activités de service aux entreprises
C75T99	75-99	Services collectifs, sociaux et personnels

## LA BASE DE DONNÉES ANALYTIQUE SUR LA DIRDE (ANBERD) DE L'OCDE : EXPLICATIONS CONCERNANT SON ÉLABORATION

### Les limites des données officielles sur la DIRDE (OFFBERD)

Le *Manuel de Frascati* donne des indications pour collecter les données de R&D, notamment sur la DIRDE, de façon à ce qu'elles soient comparables au niveau international. Bien que dans l'ensemble de la zone OCDE, les définitions et méthodologies du *Manuel de Frascati* soient globalement respectées, on note d'importantes divergences entre pays quant à la façon dont sont recueillies les données sur la DIRDE. Il en résulte d'importants effets sur la comparabilité internationale des données et la cohérence des séries chronologiques. L'objet de la présente section est de faire un tour d'horizon des principaux problèmes de comparabilité internationale que posent les données officielles sur la R-D des entreprises.

Les méthodologies d'échantillonnage des enquêtes sur la DIRDE ainsi que le champ couvert par ces enquêtes peuvent différer d'un pays à l'autre. Parfois, de nombreuses petites entreprises ne sont pas prises en compte dans l'enquête principale. Des enquêtes additionnelles peuvent être réalisées auprès des entreprises exclues, et les données peuvent être estimées comme il convient, mais ce n'est pas une pratique universelle et la manière dont l'imputation est réalisée varie d'un pays à un autre. De même, la façon de comptabiliser les établissements de recherche correspondant à des cas limites peut également varier entre pays. Lorsque les organismes et établissements de recherche publics ou semi-publics réalisent des activités appréciables de R-D industrielle, ces différences peuvent être significatives.

La pratique préconisée par le *Manuel de Frascati* est d'attribuer la DIRDE au niveau de l'entreprise. Lorsque ce principe est interprété de façon stricte, l'ensemble de la DIRDE d'une entreprise diversifiée sera attribué à la catégorie industrielle de son activité principale. Dans les cas où un nombre limité de grandes entreprises occupe une position dominante dans les dépenses de R-D de plusieurs domaines, cela peut conduire et conduit effectivement à sous-estimer la R-D associée aux activités secondaires des entreprises. La R-D est donc surestimée pour certaines branches et sous-estimée pour d'autres. Toutefois, les pays n'ont pas tous une interprétation stricte de l'attribution au niveau des entreprises des dépenses de R-D entre catégories industrielles. Certains pays ventilent la R-D de leurs entreprises diversifiées les plus importantes entre un certain nombre d'activités différentes ou bien les données sont comptabilisées par groupe de produits.

Le degré de comparabilité internationale entre pays comptabilisant par groupe de produits ou de semi-produits est vraisemblablement plus important que celui entre pays suivant strictement la méthode par entreprise. Un des avantages de données par groupe de produits est que les dépenses de R-D effectuées par une entreprise qui est active dans plus d'une branche sont réparties entre les branches respectives.

Il existe également d'importantes divergences d'un pays à l'autre dans la façon de comptabiliser la R-D exécutée par des entreprises qui appartiennent nominalement au secteur des services mais qui sont étroitement associées (pas nécessairement de façon contractuelle) à des entreprises manufacturières. Les établissements de recherche industrielle, financés dans une large mesure par les industries manufacturières auxquelles ils s'adressent, sont les exemples les plus fréquemment rencontrés. Le dernier *Manuel de Frascati* (2002) stipule que la R-D réalisée par ces établissements doit être attribuée à la catégorie : "Recherche et développement" (CITI, rév. 3.1 : 73), mais que « des informations supplémentaires aux fins d'analyse et de comparaisons internationales [...] seront rassemblées » en ventilant les données par groupes de produits. Par ailleurs, les dépenses de R-D des grandes entreprises

publiques, par exemple dans les secteurs des télécommunications et de l'énergie, sont fréquemment classées dans le secteur des services même si ces entreprises réalisent un volume considérable, et parfois majoritaire, de R-D correspondant aux industries manufacturières. Cela peut fausser les comparaisons internationales tant au niveau des agrégats qu'à celui des branches d'activités.

Il est également fréquent que les systèmes de classification soient révisés, ce qui peut entraîner des changements significatifs dans la concordance entre les classifications industrielles nationales et la classification CITI, et par conséquent, des ruptures dans les séries chronologiques correspondantes de la DIRDE. En outre, certains pays ne procèdent pas chaque année à des enquêtes sur la R-D des entreprises, d'où une couverture temporelle discontinue. Comme la DIRDE augmente rapidement dans la plupart des pays de l'OCDE d'une part, et qu'elle est très variable d'autre part, il convient de faire preuve d'une grande prudence pour établir des comparaisons internationales basées sur des chiffres correspondant à des années différentes, même lorsque celles-ci sont peu éloignées.

Enfin, les données peuvent ne pas être disponibles pour toutes les branches dans certains pays. Fréquemment, pour des raisons de confidentialité ou autres, des données pour les différentes branches ne sont pas disponibles séparément, mais associées ou combinées à des données relatives à d'autres branches. Ce type de pratique varie souvent dans le temps, de sorte que pour certaines années les industries seront disponibles individuellement, tandis que pour d'autres les chiffres ne sont pas indiqués ou sont combinés à d'autres données. Les séries chronologiques peuvent donc comporter d'importantes discontinuités qui rendent impossible toute comparaison internationale.

De ce fait, les données officielles sur la DIRDE (OFFBERD) sont constituées de séries chronologiques fragmentées qui doivent être utilisées avec prudence.

### **La base de données analytique de l'OCDE sur la DIRDE (ANBERD)**

Même accompagnées d'un ensemble détaillé de notes explicatives, les données OFFBERD sont peu utilisables pour l'analyse, en raison de leur couverture temporelle incomplète et du grand nombre de discontinuités qu'elles comportent. C'est pour répondre à cela qu'a été conçue la base de données ANBERD. Cette dernière a été construite afin de disposer d'un ensemble de données cohérent et comparable sur le plan international pour les dépenses industrielles de R-D. Cependant, de par sa nature même, ANBERD contient de nombreuses estimations de l'OCDE qui peuvent différer sensiblement des données officielles correspondantes. Du fait que ANBERD dépende d'un certain nombre d'estimations constamment affinées et réexaminées, il convient de considérer cette base comme un outil susceptible d'être révisé.

ANBERD s'appuie sur les chiffres officiels de la DIRDE, assortis d'estimations chaque fois que :

- des problèmes importants se posent du fait de l'utilisation de l'entreprise comme unité statistique ou pour la classification des cas limites. Chaque fois que cela est possible, les estimations ANBERD font l'objet d'ajustements pour faire en sorte que la DIRDE des établissements correspondant à des cas limites et des entreprises publiques soit attribuée aux branches appropriées ;

- des écarts importants sont observés par rapport à la classification industrielle type de la DIRDE (par exemple l'industrie "Construction aéronautique" est incluse dans celle des "Véhicules automobiles") ;
- des ajustements importants sont nécessaires quand le champ de l'enquête est incomplet ;
- les séries comportent des discontinuités ou des ruptures dues à une modification dans la classification industrielle ou les techniques d'enquêtes ;
- des données manquent pour des années entières lorsque les enquêtes ne sont pas réalisées chaque année.

L'ensemble de données ANBERD est construit en étroite collaboration avec les autorités statistiques nationales et approuvé pays par pays. Dans tous les cas, des techniques d'estimation statistique sont utilisées pour compléter les valeurs manquantes ou ajuster les données OFFBERD existantes afin d'obtenir des séries chronologiques améliorées et plus complètes. On trouvera dans la prochaine section une description succincte des techniques utilisées pour établir les estimations ANBERD.

### **Nouveautés depuis l'édition 2009**

La base ANBERD est maintenant actualisée de "manière glissante". Les tables sont publiées pays par pays dès leur finalisation.

Dans l'édition de 2009, les industries "Agriculture" et "Extraction" ont été incluses dans les agrégats nationaux (DIRDE totale) même si elles ne figurent pas séparément dans les tableaux détaillés ANBERD. Dans la toute dernière édition, ces deux industries sont publiées séparément. En outre, deux pays supplémentaires sont disponibles : l'Estonie et le Mexique.

## **ESTIMATIONS ANBERD**

Cette section présente l'état actuel de l'élaboration de la base de données ANBERD. Les séries chronologiques commencent en 1987, sont fondées sur la classification CITI, révision 3.1 et sont disponibles pour l'ensemble des pays de l'OCDE (l'Australie, l'Autriche, la Belgique, le Canada, le Chili, la République tchèque, le Danemark, l'Estonie, la Finlande, la France, l'Allemagne, la Grèce, la Hongrie, l'Islande, l'Irlande, Israël, l'Italie, le Japon, la Corée, le Luxembourg, le Mexique, les Pays-Bas, la Nouvelle-Zélande, la Norvège, la Pologne, le Portugal, la République slovaque, la Slovénie, l'Espagne, la Suède, la Suisse, la Turquie, le Royaume-Uni et les États-Unis) ainsi qu'une sélection de pays non membres (Afrique du Sud, Chine, Fédération de Russie, Roumanie, Singapour, Taipei chinois).

Les estimations ANBERD sont exprimées en monnaie nationale et en parité de pouvoir d'achat (\$ PPA) aux prix courants et constants. Les estimations ANBERD pour la zone Euro sont présentées en euros (EUR). Les données antérieures à l'année d'accession à la Zone Euro ont été converties en EUR en appliquant le taux de change irrévocable EUR/monnaie nationale. Cette approche permet des comparaisons à l'intérieur d'un même pays au cours du temps, et garantit que l'évolution historique soit

préservée. Cependant, les données exprimées en EUR des pays de l'UME antérieures à 1999 ne peuvent être additionnées pour générer un total de la zone Euro ou pour faire des comparaisons entre les pays.

Les séries d'indices de prix du PIB ainsi que des PPA utilisées pour convertir et déflater les données ANBERD sont disponibles dans la base des [Indicateurs Economiques pour PIST \(Principaux indicateurs de la science et de la technologie\)](#).

## **Données historiques**

Il est toujours possible de consulter l'ancienne (2009) édition (lien vers OECD.STAT : [ANBERD 2009](#)) ainsi qu'une archive de la base en CITI révision 2 ([ANBERD rév.2](#)) fournissant des données pour 15 pays de l'OCDE plus la zone « Union Européenne » de 1973 à 1997/98.

## **Méthodes générales d'estimation dans ANBERD**

Les principaux problèmes associés aux données officielles DIRDE des pays ANBERD sont les suivants :

- Des données manquantes pour certaines branches combinées à d'autres branches ou ne figurant que dans des sous-totaux.
- Des ruptures dans les séries en raison de révisions de la classification industrielle. L'inclusion ou l'exclusion soudaine de secteurs industriels sont très fréquentes. Par exemple, les "Ordinateurs et machines de bureau" qui peuvent être compris dans les "Autres machines, n.c.a." pour certaines années, mais disponibles séparément pour d'autres.
- Des données manquantes pour des années entières car certains pays ne réalisent des enquêtes de R-D que tous les deux ou trois ans. C'est notamment le cas de l'Autriche, de la Grèce, de l'Irlande (jusqu'en 2000), de la Suède et de la Fédération de Russie.
- Pour certains pays, des établissements de recherche industrielle ont été classés dans le secteur des services alors qu'ils sont étroitement associés (pas nécessairement de façon contractuelle) à des entreprises manufacturières. Cela peut fausser les comparaisons internationales.

Pour faire face à ces problèmes, on fait tout d'abord appel aux autorités statistiques nationales des pays en question, ce qui permet de résoudre nombre de problèmes posés par des secteurs pour lesquels les données sont manquantes et par des ruptures de séries. Dans un deuxième temps, lorsque cela était nécessaire, la DIRDE a été estimée en appliquant des méthodes d'interpolation.

## **Méthode d'estimation ANBERD pour les Etats-Unis**

Pour les données antérieures à 2008, environ un quart à un tiers des données ventilées par industries concernant la DIRDE intra-muros totale officielle a été supprimée pour des raisons de confidentialité liées au financement par le gouvernement fédéral. Pour pouvoir inclure les États-Unis et permettre des comparaisons internationales des résultats de la R-D, il a été nécessaire d'estimer les données pour les

industries manquantes. Ces estimations sont fondées sur la publication annuelle de la National Science Foundation (NSF) intitulée *Research and Development in Industry*, qui contient des tableaux complets sur la R-D financée par les entreprises. Ces données étaient fondées sur le recensement annuel conduit par le Bureau of the Census pour la NSF. Pour la plupart des industries, la R-D financée par les entreprises est suffisante pour estimer la R-D totale car le financement par le gouvernement fédéral ne dépasse pas 10%. Dans ANBERD, pour estimer le total de la R-D exécutée dans chaque industrie, il suffit d'estimer la composante fédérale de la DIRDE pour les industries supprimées. En ajoutant ensuite l'estimation de la R-D financée sur fonds publics aux données publiées sur les dépenses de R-D des entreprises, on obtient la DIRDE totale.

La part de la DIRDE financée par le gouvernement fédéral a été estimée. Pour certaines industries, on a supposé que cette part était très faible, de sorte qu'il était possible d'estimer avec une certaine confiance la DIRDE totale de ces industries. On a ainsi pu établir des estimations pour d'autres industries ou sous-totaux par simple addition ou soustraction. Dans d'autres industries, on a déterminé la composante fédérale de la DIRDE par interpolation linéaire entre les données pour lesquelles la ventilation entre les dépenses fédérales et les dépenses des entreprises était connue. Le résultat est une matrice complète -- pour tous les secteurs manufacturiers manquants à l'exception de la "Construction navale", ce dernier étant inclus dans "Autres matériels de transport" de la classification industrielle des États-Unis.

A partir de 1999, les données de la NSF sont publiées dans la classification industrielle nord-américaine (NAICS) remplaçant la "Standard Industrial Classification" (SIC). Même si les données restent comparables à celles des années précédentes en ce qui concerne le Total entreprises, le total des Activités de fabrication et le total du Secteur des services, des ruptures de séries peuvent apparaître à un niveau plus désagrégé de la CITI, Révision 3.1.

A partir de la fin des années 1990, un volume de plus en plus important de R-D a été attribué aux industries de commerce de gros, mais cette croissance s'est avéré être due en grande partie à deux facteurs : (1) la méthodologie fondée sur les dépenses salariales utilisée pour la classification par industries ; et (2) le changement vers la NAICS en 1999. Ceci a conduit à une révision de la méthode utilisée pour classifier la R-D par industrie à partir de l'année de référence 2004. L'impact majeur de cette révision a été une augmentation de 40% de la R-D classifiée dans le secteur manufacturier (surtout dans les produits pharmaceutiques et les TIC), surtout aux dépens de l'industrie de commerce de gros.

Après 2007, l'enquête industrielle sur la R-D (SIRD) a été remplacée par l'enquête sur l'innovation et la R-D dans l'industrie ("Business R&D and Innovation Survey") et les estimations ANBERD ont alors été basées sur la publication de la NSF "Business R&D and Innovation". Comme auparavant, certains chiffres sur la DIRDE, associés au financement de la R-D par l'Etat ou par d'autres sources, ont été supprimés pour des raisons de confidentialité. De façon à inclure les États-Unis dans les statistiques de R-D comparables au niveau international, ces industries manquantes ont été estimées.

### **Données par groupes de produits**

Suivant les recommandations du dernier *Manuel de Frascati* (2002), l'un des objectifs d'ANBERD est de présenter des données par groupes de produits lorsque cela est possible. Étant données les différentes méthodes de classification des entreprises dans les enquêtes nationales, et malgré les efforts d'harmonisation, des précautions doivent être prises lorsque l'on compare les données de certains

secteurs entre différents pays (voir aussi les notes par pays sous chaque tableau de données). Par exemple :

- Certains pays collectent et transmettent des données par groupes de produits (par exemple la Finlande, le Royaume-Uni et la Suède) ;
- Pour d'autres, des ajustements sont faits à partir de données par groupes de produits afin de réduire les biais relatifs aux données au niveau entreprise (par exemple le Japon, le Norvège et l'Espagne) ou en redistribuant la R-D de CITI R-3 73 (services de R-D) par industrie bénéficiaire ;
- Enfin, pour les pays restants, les données par groupes de produits ne sont pas disponibles, donc les données se rapportent à l'activité principale avec des estimations complémentaires de l'OCDE pour les données manquantes (par exemple le Canada et l'États-Unis) ;
- Trois pays (l'Afrique du Sud, la Belgique et la République tchèque) disposent de données selon les deux critères, d'où l'inclusion de deux tableaux dans ANBERD.